THE MEASUREMENT OF CHANGES OF THE GENERAL PRICE LEVEL

SUMMARY

Aim and method of the present inquiry, 557. — The two fundamental problems, 558. — Prices should be treated as ratios, 559. — When so treated the paradox of the arithmetic and harmonic means is resolved, 559. — Ratios are not fractional quantities, 561. — The completely weighted arithmetic mean is a ratio of aggregates, 563. — Geometrical representation of the price level, 563. — The method of least squares, 566. — The method of summation, 568. — Changes of the general price level, 568. — Effect of changes of the relative importance of trade in different types of commodities, 571. — Substitution of "dollar's worths" for physical units, 571. — Fisher's index number, 572. — Conclusion, 573.

THE violent price changes of the past few years have given new importance to index numbers and new interest to the problem of the best way of constructing them. It may be too much to hope that any general agreement upon the right solution of this problem is in sight. But substantial progress toward such agreement has unquestionably been made. The debatable issues have been narrowed. There are, for example, few who would now hold that one type of index number very commonly used in the past — the arithmetic average (weighted or unweighted) of the percentage ratios between prices in a given year and prices in a basing year, *i. e.*, the arithmetic average of "relative" prices — is other than unsound in principle and misleading in the results it gives.

Possibly the most fruitful recent work in this field has involved careful comparison of the results yielded by

different types of index numbers.[1] Agreements, differences, and anomalies in these results have been traced back to their roots. In this way much light has been thrown upon such matters as the choice of the original data, the peculiarities of different sorts of averages, the effects of weighting, and the results of choosing one basing year rather than another. But the present paper is born of the conviction that to such empirical studies, valuable as they have proved themselves to be, there must be joined a frontal attack upon the theoretical issues involved in the problem.

There can be no final agreement upon any system of index numbers until there is agreement upon the precise meaning of the phenomena they are supposed to record. We must formulate our problem before we can hope to solve it. This paper is offered, therefore, as a preliminary exploration in the field of the general theory of index numbers of price movements. In this field there are two fundamental questions: (1) Just what are the measurable phenomena defined by the phrase, *changes of the general level of prices?* (2) Similarly, what is the precise meaning to be given to the phrase, *average changes of prices?* For the present I limit myself to the first of these two problems. I hope to deal with the second in a subsequent paper.

For convenience we may begin by taking the general level of prices as defined by P in the equation of exchange, $P = M/T$, where M is the amount, in money units, of money and of bank deposits exchanged during a year or other definite short period of time for the goods whose general price level is sought.[2] T is the quantity of

[1] In this connection any student of index numbers will think of the work of Fisher, Mitchell, and Persons. The other recent writers on the problem to whom I wish to acknowledge an especial measure of personal indebtedness are Messrs. Flux and Walsh.

[2] M, as used here, corresponds to $(MV + M'V')$ of Professor Fisher's formulation of the equation of exchange.

goods exchanged for M, measured in the physical units to which price quotations refer. P is thus the (arithmetic) average amount of money paid by buyer to sellers per physical unit of goods purchased. The problem is to determine a series of numbers whose variations shall have the closest possible correlation with the variations of P in successive periods of time.

Now a price, as a statistical datum, is *always a ratio.* A quantity of money is not, for statistical purposes, a price, apart from its relation in exchange to a definite quantity of goods. Just as P, the average price of all goods bought and sold, is the ratio between M and T, so any price, p, is the ratio between m and t, where m is the number of money units paid for t units of goods.

Commonly, price quotations run in terms of the ratios between various amounts of money and single units of goods, that is, in terms of *unit prices.* The consequent of the price ratio is made unity, and price is expressed in terms of the antecedent alone. But this convenient practice should not be permitted to obscure the fact that unit prices are themselves ratios, that a price of two dollars per bushel, for example, may be expressed not only as the ratio between 2 and 1, but quite as accurately as the ratio between 4 and 2, or between 30 and 15, or as any other equal ratio.

A good example of the significance of the point that in devising index numbers prices must be treated as ratios is given by what may be termed the paradox of the arithmetic and harmonic means. "In European countries," says Professor A. C. Pigou,[1] "price is usually measured by naming the number of units of the standard of value which will buy a unit of the commodity; in India it is measured by naming the number of units of the commodity which can be purchased by a unit of cur-

[1] Wealth and Welfare, pp. 35, 36.

rency. The choice between these two ratios is obviously a pure matter of arbitrary convention. But, to combine price ratios taken on the Indian plan into an arithmetic mean is equivalent to combining similar ratios taken on the European plan into a harmonic mean!"

By "price ratios," it seems from the context, Professor Pigou means relative prices — the ratios of prices in one year to the prices of the same goods in another year. Thus if the unweighted arithmetic mean of relative prices is $\Sigma\left(\frac{p_1}{p_0}\right)/n$, the reciprocal of the harmonic mean is $\Sigma\left(\frac{p_0}{p_1}\right)/n$, which is the arithmetic mean of relative prices expressed in the Indian manner, such as $\frac{1/p_1}{1/p_0}$. The effect of using the reciprocal of the harmonic mean of relative prices instead of their arithmetic mean, it will be noted, is to shift the base from one to the other of the two years involved in the comparison. That is, the reciprocal of the unweighted harmonic mean of relative prices is identical with the unweighted arithmetic mean, except that the base is shifted.

The significance of this lies in the fact that, as is well known, when basing years are shifted the results indicated by the use of the arithmetic average of relative prices are modified. Neither the harmonic mean of a group of ratios or its reciprocal will accord with their arithmetic mean. Yet prices quoted in the Indian manner are the reciprocals of prices quoted in the way to which we are accustomed. What is more important, the market facts are wholly independent of the manner of quotation. "Five cents a pound" and "twenty pounds for a dollar" are identical prices. Where the basic facts are the same the two methods of quotation should lead to identical index numbers.

Similar considerations hold with respect to unweighted averages of simple or "absolute" prices. In general the arithmetic mean, $\Sigma(p)/n$, is not in agreement with the harmonic mean, $n/\Sigma\left(\frac{1}{p}\right)\cdot$ And no system of weighting, unless it be wholly arbitrary, will bring the two means into agreement, so long as the prices averaged are treated as quantities instead of ratios. To treat p as merely a sum of money, five cents for example, and its reciprocal $1/p$ as a quantity of a commodity, twenty pounds for example, is to handicap ourselves by a serious initial error.

But if we hold rigorously to the fundamental fact that price is always a ratio, involving two magnitudes — a quantity of money *and* a quantity of goods — the problem may be said to solve itself. The way in which prices happen to be quoted does not affect the result. Following the notation already indicated, put m/t for p. Then the only arithmetic mean which has any obvious meaning is the weighted mean, $\Sigma(m)/\Sigma(t)$, or $\Sigma(tp)/\Sigma(t)$. In a similar way the weighted harmonic mean is $\Sigma(m)/\Sigma\left(m\frac{1}{p}\right)$, or $\Sigma(tp)/\Sigma(t)$. It is identical with the weighted arithmetic mean. This weighting of the arithmetic mean of actual prices by physical quantities of goods and of the harmonic mean by quantities of money, or "values," is in no sense arbitrary. It is clearly indicated or even compelled by elementary considerations.

The paradox which has just been discussed, like a host of other difficulties which beset the construction and interpretation of index numbers, springs from erroneous methods of handling ratios. These errors have crept into the treatment of the problem, I believe, through the habit of identifying ratios with the frac-

tions commonly used to express them. Fractions whose numerators are the antecedents of ratios and whose denominators are their consequents may be subjected to certain mathematical operations without affecting the essential nature of the relations they express. But addition and subtraction are not always among these legitimate operations. Ratios are not always safely treated as additive quantities.

Moreover, when the fractions used to express ratios are added as part of the process of finding their arithmetic mean, they are first reduced to a common denominator. This involves an arbitrary weighting of the numerators. In common practice the numerators are determined by the condition that the denominator shall in each case be unity. Then the denominators are forgotten, the numerators are termed "prices," and their mean is held to be an "average price." What the unweighted arithmetic mean of actual prices really gives is merely the average money cost per unit of a bill of goods consisting of one unit each of the different commodities whose price quotations are used. Similarly the reciprocal of the unweighted harmonic mean gives the average number of physical units of goods exchanged for one unit of money. The difference between the results given by the two means is due to the presence in each case of weighting that is none the less arbitrary because it is not intended. In one case single units of goods and their money costs are the components of the average. In the other case its components are units of money (dollars) and "dollar's worths" of goods. One mean has as good a claim to represent the general price level as the other. But for neither is the claim valid unless it is expressly understood that the mean represents merely a group of prices whose peculiar composition is dictated by the method by which the mean is constructed.

We have seen, however, that with *complete weighting* the two means lead to consistent results. The reciprocal of the arithmetic mean becomes identical with the arithmetic mean of the reciprocals of the components of the original mean. This illustrates an elementary but important difference between the unweighted and the completely weighted arithmetic mean of ratios. The unweighted arithmetic average of ratios is a "fictitious mean" of the familiar type, but its claim to represent or typify a group of ratios is marred by its hidden weighting and especially by the fact that it gives different results according as the ratios whose mean it purports to be are expressed by one form of fraction or another. But the completely weighted arithmetic average is not a "fictitious mean." In fact its title to the name average or mean might even be disputed. It is nothing but a *ratio of aggregates.* It is constructed by a process of summation rather than of "averaging." Its antecedent and consequent are the respective sums of the antecedents and consequents of its constituent ratios.

Consider an example taken from another statistical field. The proportion of the population of each state found in some specified population group, those for example who are gainfully employed, is reported by the federal census. For any one state this proportion is the reciprocal of the ratio of the whole population of the state to the number who are gainfully employed. An unweighted mean of the per cents which express the proportions found in the different states might conceivably be held to represent the proportion gainfully employed in an imaginary "average state." But its reciprocal is likely to differ, perhaps rather widely, from the unweighted mean of fractions which have whole populations as numerators and numbers gainfully employed as denominators. Yet this second mean has

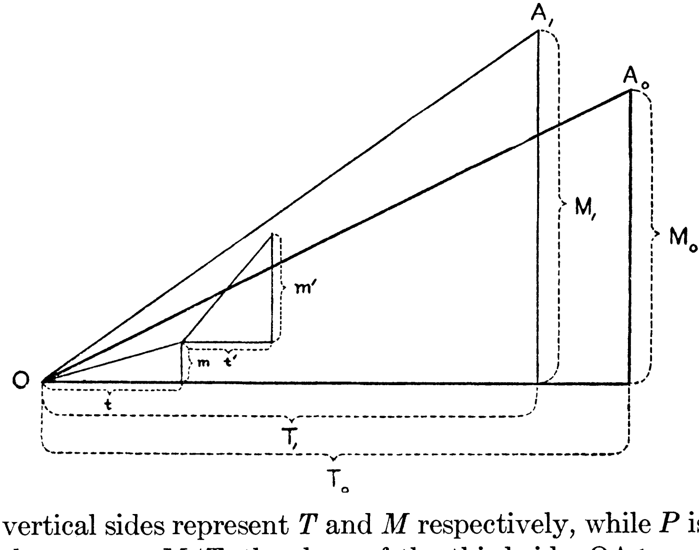
as valid a title as the first to be the ratio attributed to the "average state."[1] With complete weighting the two means are brought into agreement. But they no longer indicate a "fictitious," "representative," or "typical" condition. They express for the United States as a whole the ratio between the aggregate number gainfully employed and the aggregate population. In a similar way the completely weighted arithmetic mean which we have been discussing is not, properly speaking, an average of separate classes of prices. The ratio between the total amount of money exchanged for goods and the total quantity of goods exchanged for money directly expresses the general or aggregate price level.

It should be clear, of course, that this method of summation, of finding a ratio between aggregates, may be used appropriately in obtaining a general expression for other groupings of price ratios than those which correspond to P in the equation of exchange. The general price level of goods produced, of consumers' goods, of exports or imports, etc., may likewise be expressed as ratios of aggregates. For each special problem there is a different selection and a different weighting of the price ratios which enter into the aggregates. The ratio P, however, alone has a right to be termed *the* general price level. For price is a fact of exchange rather than of production or consumption. If prices are themselves to be taken as our fundamental facts, all the prices that

[1] A somewhat similar problem is created by the requirement of the Constitution that "Representatives shall be apportioned among the several states according to their respective numbers." The arithmetic mean will in general lead to different apportionments according as "representatives" or "respective numbers" are denoted by the numerators of the fractions used to express apportionment ratios. Precedent and tradition support the present practice of using the first of these two forms. This gives results which in the long run are more favorable to the larger states than would be obtained by using the reciprocals of such fractions. Dr. J. A. Hill and Professor E. V. Huntington have shown that this particular problem calls for the use of the geometric rather than the arithmetic mean. With the geometric mean, of course, the form the fractions take is a matter of indifference, for the geometric mean of the reciprocals of fractions is the reciprocal of the geometric mean of the fractions in their original form.

occur have to be taken into account. That is, all exchanges must be taken into account. And since exchange involves two sorts of magnitudes — quantities of money and quantities of goods — neither magnitude taken alone gives a safe basis for weighting.

Some aspects of the matter may be brought into a clearer light by a geometrical device. The equation of exchange may be represented by a right-angled triangle, as in the accompanying diagram. The horizontal and



vertical sides represent T and M respectively, while P is the tangent M/T, the slope of the third side, OA.[1]

Knowledge of the absolute magnitudes of M and T would give us the value of P. But the most available and certainly the most trustworthy data are observations on the price ratios at which different goods are

[1] Similarly, the tangent M_1/T_1 is P_1, the price level in some other year. Individual weighted prices are represented by m/t, m^1/t^1, etc. They are placed on the diagram so as to illustrate the summational method, viewed as the addition of vectors. To illustrate the method of least squares these small triangles should be so placed as to have a common origin at O_1.

exchanged for money. These may be supplemented, we must assume, by some knowledge of the relative magnitudes of the volume of trade in the same goods. That is, our data are certain p's and t's. Our facts are incomplete.

What is the best method of inferring from our partia observations the most probable value of P, that is, the most probable slope of the line OA?

If each reported price (itself to be interpreted graphically as a tangent or slope) be taken as a separate and independent observation on the value of P, that is, on the tangent or slope M/T, the problem is obviously that of fitting a straight line to a number of observed points, each subject to unknown but presumably unbiased errors. The method of least squares suggests itself. This method, it will be remembered, involves the principle of the arithmetic mean. The appropriate formula is:

$$P = \Sigma\,(tm)\,/\,\Sigma\,(t^2) = \Sigma\,(t^2p)\,/\,\Sigma\,(t^2).$$

Where only one physical unit of each commodity is taken into account this formula becomes $\Sigma(p)/n$, the unweighted arithmetic mean of unit prices. Graphically this may be interpreted by conceiving the slope of the line as determined by passing it through the origin and a point fixed by the arithmetic mean of the n values of the ordinate m at the point where the abscissa t is unity.

Taking the different magnitudes of t into account, the method gives a less familiar result. The ratio $\Sigma(t^2p)/\Sigma(t^2)$ is an arithmetic mean weighted in accordance with the squares of the numbers of commodity units sold at the stated prices. That is, it is what is commonly deemed a properly weighted arithmetic mean, operated upon by again multiplying its different component terms by the weights already used.

But the old difficulty remains. Invert the triangle, using its base as an axis, and then rotate it in a clockwise

direction so that its vertical side becomes horizontal. The slope of the hypothenuse is now T/M, or $1/P$. The formula for determining this slope, according to the method of least squares, is $\Sigma(mt)/\Sigma(m^2)$, or $\Sigma\left(\frac{m^2}{p}\right)/\Sigma(m^2)$. Letting m equal unity this gives $\Sigma\left(\frac{1}{p}\right)/n$, the reciprocal of the unweighted harmonic, not of the unweighted arithmetic mean. Nor is the weighted form, $\Sigma(mt)/\Sigma(m^2)$, the reciprocal of the correlative form previously found, $\Sigma(tm)/\Sigma(t^2)$. Their product is $[\Sigma(mt)]^2/[\Sigma(m^2) \times \Sigma(t^2)]$, which is not, in general, equal to unity.

As between these two weighted forms there seems to be no ground of preference. But they may be combined or averaged so as to give each of them equal weight. This is best done by using the geometric mean. The result is $\sqrt{\Sigma(m^2)/\Sigma(t^2)}$, or $\sqrt{\Sigma(t^2p^2)/\Sigma(t^2)}$. The method of least squares, therefore, does not lead to the ratio of aggregates, but to another weighted arithmetic mean, the square root of a weighted average of squares in which the weights are also squared.

The use of this formula, novel as its application to price statistics is, might perhaps be justified in cases where the price quotations available could properly be regarded as merely a relatively small random sample of the complete series of actual prices. At any rate, this weighted average of squares is the only defensible form of arithmetic mean of actual prices which can be regarded as "representative" rather than "summational" or "aggregatic." [1]

In general, however, the analogy between the relation of observed prices to the general price level and the rela-

[1] A useful descriptive adjective, suggested by Professor Irving Fisher.

tion of points given by experimentation or observation to the equation of the best-fitting curve or line may easily be stressed too heavily. Price statistics are not, in a strict sense, independent observations of an unknown magnitude — the general price level. Each recorded observation is a partial rather than an independent report on that magnitude. The antecedent and the consequent of the reported price ratio enter into the sums which constitute the antecedent and consequent of the general price ratio.

If all exchanges were reported, the equation of exchange could be written $P = \Sigma(m)/\Sigma(t)$. And if, tho not complete, our list of reported prices is reasonably large and representative, it may be used to give a *partial sum*. That is, we take $\Sigma(m)/\Sigma(t)$ as giving a workable approximation to M/T. We must assume, therefore, either that the remainders $M-\Sigma(m)$ and $T-\Sigma(t)$ are relatively small as compared with $\Sigma(m)$ and $\Sigma(t)$, or that the unknown ratio $[M-\Sigma(m)]/[T-\Sigma(t)]$ is not likely to be greatly at variance with the ratio $\Sigma(m)/\Sigma(t)$, and as likely to be above it as below it.

In practice, as we have already seen, this method means the use of the ratio of aggregates. Graphically, this amounts to determining the slope of the line OA by passing it through a point whose ordinate is $\Sigma(m)$ and whose abscissa is $\Sigma(t)$. This is equivalent, of course, to the addition of a series of vectors, each of which is determined by a particular weighted price ratio, m/t, just as OA is determined by M/T.

With the general price level determined there is no technical difficulty in the way of comparing one year's prices with another's by constructing a series of index numbers representing the fluctuations of the general price level. It is necessary only that the price level in

each successive year be expressed as an independent magnitude. With this condition met relative changes in the price level are easily determined. The results do not depend upon the selection of a particular basing year or period.

With constant weighting $\Sigma(t)$ is constant, and the index numbers express merely the fluctuations of $\Sigma(m)$. In this form index numbers of the general price level are the "weighted sums" which several recent writers have ranked among the better index numbers, and which have proved themselves serviceable in practice. But tho free from some of the defects of other familiar types of index numbers, these weighted sums are not wholly satisfactory expressions of the general price level. They indicate merely the varying amounts of money payments required by sales of constant quantities of goods at prices determined in part by the condition that such sales have not in fact been constant. Or with other weights than t, they may indicate the changing market value of a constant volume of production, or the changing cost of a fixed bill of goods or of a fixed standard of living.

Another sort of weighted sum is possible. If $\Sigma(m)$ and its constituent parts are treated as constants $\Sigma(t)$ would serve as a weighted sum. Its variations would show the fluctuations in the volume of goods that would be sold under the condition that constant amounts of money are assigned to the purchase of each commodity at prices determined in part by the condition that money payments are not, in fact, constant.

The two different types of weighted sums give different, and inconsistent, index numbers. There is no reason to prefer one to the other as an index of the "general level of prices." For this purpose neither is satisfactory. Each has its own significance, but this

significance lies in the particular meaning determined by the method of construction. Paucity of data makes the second type impracticable except for the study of limited fields of production, consumption, and trade. But it has a better right than the other to be considered an index of the "purchasing power of money."

As compared with the weighted sum, the ratio of aggregates has not met with equal favor. This may be accounted for by a variety of circumstances. Doubtless the limitations in the meanings of other types of index numbers, including weighted sums, have not always been fully realized. Moreover, the magnitudes, even the relative magnitudes, of t (as of other possible weights that might represent physical quantities) are so incompletely and inaccurately known as to discourage attempts to estimate and utilize their fluctuations in index numbers of the general level of prices. Most of all, I suspect, it has been felt that an average into which enter such heterogeneous things as tons, yards, gallons, and dozens, is devoid of meaning.

With respect to this last objection it should be observed that for ascertaining the price level there would be no advantage in converting these incongruous physical units into some common measure, tons for example, as is sometimes done in statistics of foreign trade. This would merely introduce a new and deceptive system of weighting. Moreover, price ratios between homogeneous money units and heterogeneous units of goods are themselves comparable things, and such ratios are all that is involved in the ratio of aggregates. The price ratio itself is taken as the fundamental statistical datum. In that ratio each physical unit figures merely as the quantum in terms of which prices are made and recorded. Viewed as quantities the different sorts of physical units are, admittedly, not properly comparable one with another.

But they are used here merely as terms in ratios, and the comparability of such ratios is implied in the very phrase, "the general level of prices."

A serious difficulty appears, however, when the changing volume of trade in different types of commodities is taken into account. Other things being equal, if sales of commodities whose unit prices are low come to make a larger proportion of the total volume of exchanges, the fact will be reflected in a lower general level of prices. Similarly, if the sales of commodities whose unit prices are high grow faster than other sales, the general level of prices will be higher. While not without their own significance, changes of this kind obscure the action of larger forces.

For this reason, and, so far as I now see, for this reason only, there is some advantage in measuring trade in "dollar's worths" rather than in single physical units. Conversion of physical units into "dollar's worths" may be effected on the basis of the prices that prevail in either of any two years whose price levels are to be compared. For that year the general price level becomes unity. For the other year it becomes $\Sigma(m_1)/\Sigma(t_1p_0)$, or $\Sigma(t_1p_1)/\Sigma(t_1p_0)$, where t still represents the number of physical units sold at the unit price p, and the subscripts 0 and 1 indicate respectively the basing year and the other year.

This result, $\Sigma(t_1p_1)/\Sigma(t_1p_0)$, may be taken as an expression of the *relative price level.* Like the absolute price level it is a ratio between the volume of money payments and the quantity of goods purchased, but in this case the quantity of goods is measured not in tons, yards, bushels, etc., but in the amounts that were exchanged for a dollar in the basing year. In form, however, this result is merely a ratio between weighted

sums of unit prices, the constant weights, it will be noted, being not those of the base year, but those of the year whose price level, relative to that of the basing year, is sought. But it has a meaning and significance not usually attached to ratios between weighted sums.

One difficulty remains. If the basing year is shifted, so that the price level in what had been the basing year is now expressed as relative to that of the other year (which now becomes unity), we have the form, $\Sigma(t_0p_0)/\Sigma(t_0p_1)$. This is not in general the reciprocal of the other form, $\Sigma(t_1p_1)/\Sigma(t_1p_0)$. Here again the geometric mean may properly be used to combine the two inconsistent results.[1] We then have as the expression of the relative general price level in the year denoted by the subscript 1: $\sqrt{\frac{\Sigma(t_1p_1)}{\Sigma(t_1p_0)} \times \frac{\Sigma(t_0p_1)}{\Sigma(t_0p_0)}}$.

I have the more confidence that this is an acceptable solution of the particular problem in hand because Professor Irving Fisher has already reached precisely this result and has adjudged it to be the best of the different formulas he has examined and tested.[2] Furthermore, Professor Fisher's result has the weighty approval of Mr. C. M. Walsh.[3] So far as I know, however, the analysis by which I have reached this conclusion runs along lines which differ from those Professor Fisher and Mr. Walsh have followed. It is possible, therefore, that in attempting to find my way into the fundamentals of the problem I have done something to strengthen the logical foundations of Professor Fisher's index number.[4]

[1] The propriety of using the geometric mean hinges on the fact that if the two expressions were in agreement their *product* would be unity.

[2] Quarterly Publications of the American Statistical Association, March, 1921, p. 536.

[3] Ibid., p. 539.

[4] It is fair to say that most of the results of the present paper were presented to the Economics Seminar at Cornell University nearly two years ago. I did not at that time, however, suggest that the geometric mean be used to get a synthesis of two inconsistent results.

On the basis of the evidence now in hand I believe this to be the best single index number of the general level of prices. It is less likely to be deceptive than any other formula that I happen to know. I mean that it gives a more direct and unequivocal answer to the particular questions most students of changes of the price level are likely to have in mind. But I fear that no single index number will afford a sufficient answer to all such questions.

It should be clear, moreover, that in the present paper we have been concerned solely with the problem of changes of the general price level. The measurement of average changes in prices is in some respects a distinct and different problem.

ALLYN A. YOUNG.

HARVARD UNIVERSITY.